DOCUMENT RESUME

ED 396 369 EA 027 588

AUTHOR de Vos, Henny
TITLE Using Simulation To Study School Effectiveness.
PUB DATE 95
NOTE 16p.; Paper presented at the Annual Meeting of the European Council on Educational Research (Bath, England, September 14-17, 1995).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; Educational Policy; *Effective Schools Research; Foreign Countries; Longitudinal Studies; Models; *School Effectiveness; Secondary Education; *Simulation; Statistical Data; Statistical Inference; Student Placement
IDENTIFIERS *Netherlands

ABSTRACT

School-effectiveness research is constrained by ambiguous factors of effectiveness and a lack of theory. This paper presents findings of a study that used simulation to improve school-effectiveness theory. Simulation is also used to explore the direct effects of schools on individual learning. After introducing simulation models, the paper describes a multilevel simulation model to simulate learning in a classroom environment over several years. Three experiments were conducted to validate data structures, compare differences between schools and classes, and generate hypothetical effects of policy changes. The first experiment analyzed longitudinal data from about 4,100 students and found a correlation among student-background characteristics (socioeconomic status, gender, IQ, and achievement). The simulated data were structured similarly to the actual-education data. In the second experiment, the model was able to generate differences between schools and classes. Finally, the third experiment generated hypothetical effects for two major policy innovations. Five figures are included. The appendices contain information on secondary education in The Netherlands and a mathematical formulation of the model. (Contains 21 references.) (LMI)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 396 369

USING SIMULATION TO STUDY SCHOOL EFFECTIVENESS

Henny de Vos

University of Twente; Department of Educational Science and Technology;
P.O. Box 217; 7500 AE Enschede; The Netherlands
tel. (31)53893831 fax (31)53340891
e-mail: vosh@edte.utwente.nl

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Paper presented at ECER 1995, 14-17 September 1995 Bath

Do not cite without permission of the author

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
H. de Vos
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

ABSTRACT

From school effectiveness research it shows that effective factors can not easily be determined in school effectiveness research, are often ambiguous, and there is a lack of theory. With these results as a background simulation is used in this paper as an aid to improve theory. Furthermore simulation is used to explore the direct effects of schools on individual learning. After an introduction in simulation models, a multi-level simulation model is described to simulate learning in a classroom environment for a sequence of years. With this model three experiments are executed to validate data structures, to study differences between schools and classes and to generate hypothetical effects of policy changes.

1 INTRODUCTION

Findings from School Effectiveness Research

The most common characterization of school effectiveness research is that it relates school variables to output data. In most studies these output data consist of test scores on home language, foreign languages or mathematics. The choice of school variables depends on the approach of the researcher toward school effectiveness. Some examples of school variables are budgets and school sizes in a more economic approach and process variables like school leadership and aspects of instruction in the effective schooling approach.

On average the between schools variance found in school effectiveness research is about 11% of the total variance in student outcomes, that is schools account for 11% of the variation in student achievment (Scheerens, 1992). However the variables responsible for the differences between schools are not easily determined and often ambiguous. This applies even more to Dutch school effectiveness research than to Anglo-Saxon studies, as there are a lot of variables and few research units in the Netherlands.

In Anglo-Saxon school effectiveness research there are some factors which are repeatedly found to explain differences between schools, which are *educational leadership*, *performance oriented*, *save climate*, *high expectations*, and *frequent evaluation of student progress*.
In research of the effectiveness of Dutch schools, however, these factors are only weakly related with student achievement and correlates are not consistent. Sometimes even negative correlates are found for 'educational leadership'. A possible explanation is that the Dutch school leader probably has a more amicable relation with the teachers and that evaluation in the Dutch schools is only weakly developed.

The basic outcome of school effectiveness research is that schools matter, i.e. schools differ in their effects on pupil achievement. Although it might be argued that school effects are relatively small

BEST COPY AVAILABLE

EA 027 588

compared to family and individual effects, they are nevertheless relevant in terms of life chances of pupils. However, the field of educational effectiveness is plagued by two problems. First, there is a lack of theory and theory-driven conceptualization due to a strongly inductivist and empiricist research tradition. Secondly, there is a gap between the complexity of the available conceptual causal models and available techniques of statistical analysis (Scheerens, 1992)

Using an existing simulation model (Bosker & Guldemond, 1994) this paper focuses on the possibilities to use simulation models of school learning to explore the relation between schools and individual learning. Instead of explaining why differences occur in individual learning, this paper aims at describing how differences in learning arise. Hence we aim at getting more insight into how school influence learning and which effects can be expected when policy changes are implemented.

Overview

Section 2 of this paper will give an introduction into modelling and simulation in educational environments. In section 3 a two-level model where students are nested in classes is described. This model is used for experiments to validate outcomes and investigate whether the model simulates realistic results. The design of the experiments is explained in section 4 and the results are given in section 5. Finally section 6 discusses the results.

2 ASPECTS OF MODELLING

A general definition of a model is that it is some representation of an object or system, which can be used to answer certain questions regarding this system.

Models of Educational Processes

A lot of models and theories have been proposed that have relevance for education. However, as Snow (1973) remarked, "*Even a superficial scanning of the literature shows amazing diversity both in the use of the terms 'theory' and 'model' and in the nature of the formulations so identified* " (Snow, 1973, p.106; Haertel, Walberg & Weinstein, 1980). Haertel c.s (1980) discuss some of these theories or models of student learning in classroom settings with achievement-related outcome variables (p.e. Carroll, 1963; Cooley & Leinhardt, 1975; Bloom, 1976; Harnischfeger & Wiley, 1979; Bennett, 1978). Most of these kind of theories or models are simply frameworks for the design and organization of measuring educational processes. The models provide minimum guidance as to how many or what type of measures or indexes would yield useful information (Leinhardt, 1980).
The components of the initial conditions and instructional processes distinguished in these models are supposed to explain differences in student learning by differences in school variables or some other variable included. Since these models have a more explanatory character, the models cannot be used to describe learning.

What makes it hard to formulate models of educational systems is the complex environment. Geurts (1983) describes this complexity in four domains:

1. multi-variate complexity: phenomena are caused by many causal forces at the same time;
2. multi-level complexity: individuals, groups and other entities constitute larger social connections and influence each other;
3. multi-relational complexity: relations between parts of a system can take on various forms, like unidirectional causation, reciprocal causation, and feedback-loops, and;
4. time complexity: systems develop over time and do not have equal patterns of behaviour at various points in time.

Especially for complex systems (like educational ones) formulating mathematical models have their use in structuring knowledge, integration of ideas and views, and improvement on communication.

BEST COPY AVAILABLE

Mathematical Models

A model is called a mathematical model when the original system is described by a set of mathematical equations. Some advantages of using mathematical models instead of other problem solving methods are described in Geurts & Vennix (1989):

1. mathematical formulation forces the modelbuilder(-s) to explicitly formulate assumptions and relations
2. by using mathematical formalisation (logical) contradictions are avoided
3. detailed mathematical formulation makes deficiencies in information clear
4. mathematical models can be used for calculations on a computer (fast and at low cost):
   - consequences of assumptions are easily determined
   - sensitive spots are quickly found
   - the effects of several scenarios can be efficiently investigated

A mathematical modelling approach for complex social problems is system dynamics or socio-cybernetics. A well-known model is that of the Club of Rome (Forrester, 1971), where dramatic forecasts were made of the growth of world population, available food and the use of resources. The philosophy behind system dynamics is that it is not the precise nature of the relations that matters, but rather the network of relations.

An example of a simulation model of learning is the model developed by Levin and Roberts (1976). This model combines research on self-esteem, standards, expectations, student ability and instructions. The model contains 2 feed-back loops: one for the student adjusting his goals and efforts and the other for the teacher adjusting his expectations for the student and giving some help depending on the performance level of the student. The model consists of, among others, lagged relations, meaning that it takes some time to adjust a variable, p.e. teachers expectations. Figure 1 contains a diagram of student-teacher interaction as it affects classroom performance.

The authors summarize the model as follows (Levin & Roberts, 1976, 99):
"In review, during the course of a school year, a child adjusts his goals so they are in line with his performance. The gap between how well a child is performing and how he would like to perform determines the amount of teacher help the child thinks he needs. This in turn influences the amount of help the student seeks, which affects the amount of time and help the teacher gives the student. This help, combined with the student's innate potential and current store of knowledge, influences his rate of learning. Learning rate increases as the student's knowledge base increases. This knowledge accumulation will determine the student's performance."

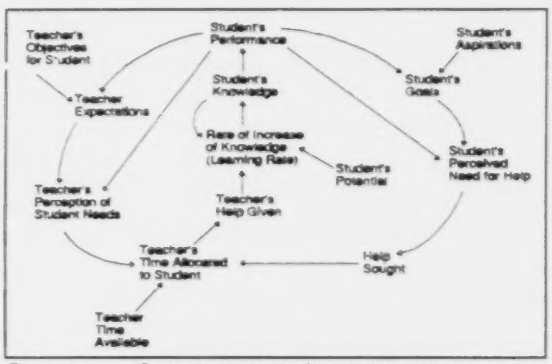

Figure 1 Systems diagram of student-teacher interaction. (Source: Levin & Roberts, 1976)

Because of its one-level nature, a model like the one of Levin and Roberts can be used to describe learning of one student, who's behaviour is influenced by the teacher. Of course the teacher has to divide his attention among the many students in his class, but this doesn't affect individual learning directly.

One-Level and Multi-Level Models

The student-teacher interaction model of Levin and Roberts distinguishes only one-level of education. As such, it ignores the fact that students are nested in classes and classes are nested in schools. Even more levels can be distinguished when one takes into account that schools are part of a broader educational system.
Hence we prefer models of a multi-level nature as these are more realistic and incorporate grouping effects. The model introduced in the next section, describes how student learn in a group.

3 A TWO-LEVEL MODEL WITH STUDENT-CLASSROOM INTERACTION

The simulation model described in this section was developed by Bosker & Guldemond (1994) for the purpose of exploring the effects of implementation of two major policy programmes to improve the functioning of secondary education (Common Core Curriculum and Educational Priority Programme). It was assumed that these innovations would raise classroom standards in the lower tracks and would discriminate in allocation of teachers time between lower and higher achieving students. Hence classroom standards and teacher time are the central variables in policy instruments in the simulation model.

The model has a hierarchical structure, in the sense that two levels are incorporated in the model: student level and class level, thus reflecting the nesting of pupils in classes. Another feature is that the model simulates student learning over time, hence the simulation has longitudinal characteristics and is able to represent the dependency of adjacent grades by those pupils who repeat a grade and the dependency of adjacent curricular tracks by those pupils who drop out of one track to proceed in the next lower one.

The Model

Assumptions and Restrictions

The model is restricted to study a *categorical* educational system with four curricular tracks with varying degrees of difficulty. Furthermore this system is bounded to *four grades*. These restrictions are a simplification of the Dutch educational system in which most schools are comprehensive and where the period of education varies from 4 to 6 years (see Appendix I for a description of Dutch education on the positioning of the simulation).
Furthermore it is assumed that achievement is one-dimensional, which means that achievement is measured with just one value. This of course ignores the fact that within secondary education many subjects are taught. In this respect one might consider the overall achievement variable in the model as an average indication of student performance.

Model Description

The model aims at simulating student learning within a school environment. Students flow to a track of secondary education from primary education. This track is not simulated, but is known at the start of the simulation since the choice for the track depends on the advice of the teacher and school leader of the primary school, preferences of parents and results of the performance tests. The simulation model has a multi-level structure, meaning that some of students following the same track and grade form a class. This multi-level structure represents the nesting of students in classes, where student behaviour is influenced by their fellow-students.
This class is educated by a teacher, who has some characteristics regarding effective instruction and time available and who sets some standard for the class. Depending on the initial achievement level, background, the teacher and the composition of class, each individual student increases (or sometimes decreases) his achievement level, i.e. the student learns. After one simulated year of education the teacher decides whether the student moves up, switches tracks or will repeat the

grade. This decision depends on students performance level relative to the performance of their fellow-students.
The procedure described is repeated for every class. This means a class can consists of repeaters, students from other tracks and students from the previous grade.

The variables influencing learning are related to an individual student, the student group and the teacher. Some variables are assumed to have a direct effect on learning (p.e. teachers time), whereas others have only an indirect effect. The relational structure of the variables is visualized in Figure 2.
A class is filled with pupils from either primary education, pupils moving up, repeating the grade or switching tracks, all requiring the same education (track and grade).
Learning is influenced by classroom characteristics, teacher variables and individual variables. At the end of the learning period (a year) a track placement decision is made for each student and it is decided whether the student leaves school, moves up, switches tracks or repeat the grade. This procedure is repeated until all students finish their education. It should be noted that the simulation processes one class at the time, but a student moving up is set in a class with the classmates of the previous grade who also moved up.

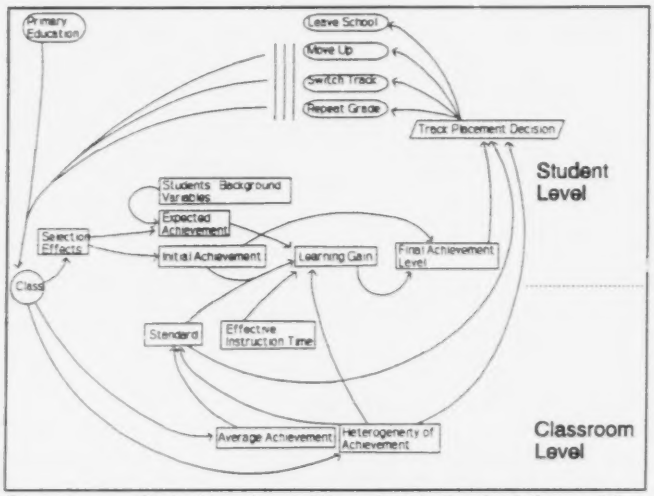

Figure 2 Student learning and throughput in the simulation model

How learning is determined

The relations, depicted by the arrows in Figure 2, are quantified using statistical and mathematical formulations. The main factors determing learning gain are teacher time available (symbol: T_{ij}), the standard set by the teacher and initial achievement level of the students, and heterogeneity of the class (symbol: σ). The precise formulation of the relations between the variables in the model is outlined in appendix II. The dependency of learning gain on these variables is illustrated in Figure 3.
The horizontal axes shows how students' achievement level is related to the average achievement of the class and the standard. Classroom heterogeneity (σ) is depicted on the vertical axes. The learning curves are drawn for teacher time available, T_{ij}, of 10, 55 and 100.
The marginal decreasing effect of time is indicated by the diminishing distance between the curves, as for instance the difference between the line T_{ij} = 10 and T_{ij} = 55 is less than the distance from the curve of T_{ij} = 55 to T_{ij} = 100. This means an increase in time from 10 to 20 has a greater impact on learning than a rise from 80 to 90.

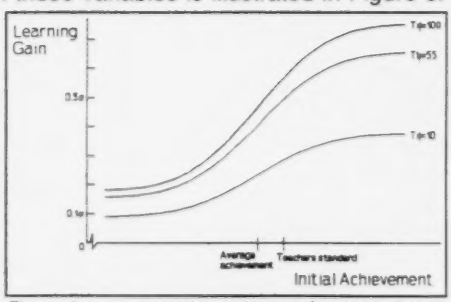

Figure 3 Learning gain as a function of time (T_{ij}), heterogeneity (σ) and classroom characteristics

The direct and indirect effects of heterogeneity on learning

From Figure 2 it can be derived that the class level variable heterogeneity has a large impact on achievement and track placement, since heterogeneity influences 3 variables. Heterogeneity has a direct effect on learning gain and influences learning indirectly via the standard. Moreover heterogeneity influences the track placement decision. The measure for heterogeneity is standard deviation of achievement. The relations are quantified in such a way that students in heterogeneous classes can make more learning gain than students in homogeneous classes. Furthermore students in heterogeneous classes face higher standards, but lower critical level for moving up.

For example consider a student Peter achieving at level 92. Peter is an underachiever in a class with an average achievement level of 100. When Peter's class is homogeneous, say a standard deviation of 5, Peter will never be able to move up, because he can never reach the level needed. However when Peter sits in a very heterogeneous class, say with a standard deviation of 15, he has a probability of 94% to move up (depending on teacher time available and random correction). The dependency of learning and track placement on heterogeneity is illustrated in Figure 4.7

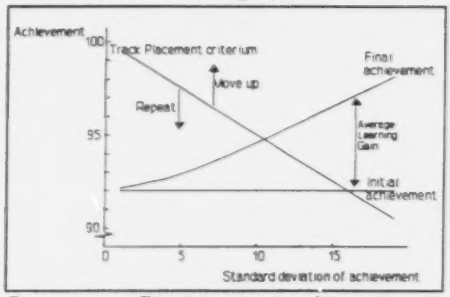

Figure 4 The dependencies of learning and track placement on heterogeneity

4 DESIGN OF EXPERIMENTS

Student data from a national cohort study are used as an input for the simulation. Background variables of the students are SES, sex, IQ and teachers rating. SES measures are reduced to a 3-point scale (low, moderate and high). It is obvious that sex is measured on a 2-point scale, where 0=male and 1=female. At the end of primary education the teacher gives each pupil an advice for secondary education. This advice or rating ranges from 1=very low (for individual types of education) to 13=very high (for pre-university level). This advice is subjective an based on teachers perspective of student performance. The fourth student variable is IQ. This is measured by a test in the last grade of primary education.

Three experiments with the simulation model will be described in the next section. The purpose of the experiments is to investigate whether the model generates realistic outcomes. The first experiment aims at confronting the data structure of simulated data with the data structure of a national cohort. A second experiment is used to investigate the differences between schools in the simulation and how these differences can be explained. In the third and last experiment hypothetical effects of some policy changes are generated.

5 RESULTS

Experiment I Validation of Data Structure

In research on student learning often pupil's achievement is corrected for student background characteristics, such as SES and gender. In other words student background has some effect on student learning and hence on school performance. Whether these relations can also be found within the simulation is studied in this section. The experiment results in a comparison of the structure of student data within the simulation and student test data of a Dutch national cohort study.

The validation experiment focuses on the correlational structure of achievement and student background variables (SES, IQ, rating, sex). Because the model does not pretend to forecast individual achievement of students, comparing individual outcomes is not desired. The simulation

BEST COPY AVAILABLE

model rather aims at generating hypothetical effects of school policies. Hence the general structure of the results is of importance.

Technical aspects

Input for the simulation are longitudinal (1989-1991) data from about 4100 students. Before entering secondary education in 1989, achievement level and IQ of these students measured by some tests. Also SES, gender and teachers rating for each student were known. In 1991, after two years of secondary education about 2300 of the students were tested again on mathematics and language knowledge. Correlations between achievement and student background were calculated for each particular track, since we assume that there might be differences between tracks.

Results

The results of one simulation are presented, as for several simulation runs a similar correlational structure was found. The correlational structure of the data for each track is given in Table 1.

Table 1 Correlation of simulated achievement and test scores with student background variables

	JVE			IGSE			HGSE			PUE		
	Math	NL	SA	Math	NL	SA	Math	NL	SA	Math	NL	SA
SES	.08	-.01	.02	-.02	.04	.04	.08	.11*	.02	.09	.07	.06
Rating	.26*	.11*	.15*	-.10*	.14*	.07*	.13*	.03	.16*	.16*	.20*	.21*
Sex	.06	.01	-.13*	.04	.03	-.04	-.10	.08	-.05	-.09	-.04	-.11*
IQ	.06	.10	.16*	.14	.11*	.10*	.05	.05	.04	.14*	.10	.06
Test results at end of PE	.12*	.44*	.73*	.01	.28*	.57*	.11*	.26*	.69*	.39*	.35*	.62*
Number of cases	534	534	752	858	858	1388	449	449	748	511	511	793

1-tailed significance -.01 indicated by *

Math	Result on mathematics test
NL	Result on language test
SA	Simulated achievement

JVE	Junior Vocational Education
IGSE	Intermediate General Secondary Education
HGSE	Higher General Secondary Education
PUE	Pre University Education
PE	Primary Education

Although correlations are low, there is some variation in correlations between the tracks.
Except for the language test in Higher General Secondary Education, SES does not seem to be correlated with achievement since correlations are not significant, for both cohort data and simulation data. Between rating and achievement positive correlates are found for JVE, HGSE and PUE. Furthermore in these tracks real correlations and correlates with simulated achievement are of the same sign and about the same level.
The third correlate, of gender and achievement, is only significant for simulated achievement in JVE and PUE.
Correlations between IQ and achievement are all positive although not every correlation is significant. Here correlations between tests and simulated achievement are almost similar.
Finally the correlation was calculated for the achievement level of 2 years ago, i.e. the achievement level at the end of primary education. As the simulation takes this achievement as a basis, correlation with simulated achievement is relatively high (about 60-70%). Here correlations for mathematics tests differ a lot between the tracks: from about 10% in JVE and HGSE to about 40% in PUE.

From these results it can be concluded that the simulated student data more or less has the same structure as data of real education. Most correlations show the same sign and about the same level for simulation data and cohort data.

Experiment II Differences Between Schools

This experiment describes the variation in student learning. An average of about 11% is found in school effectiveness research for differences in school explaining the variation in student performance.

The major question to be answered is whether the simulation model is able to distinguish between schools and how much of the variance in student learning can be explained by school of classroom level variables. In this respect it is desirable that variation in simulated student learning due to student grouping is comparable with results found in empirical research.

Technical aspects

First of all it must be noted that differences between classes reflect differences between schools as the simulation only distinguishes between classes. For the analyses of between schools variance, performance results are used of students from the first and fourth grade. As an operationalization of learning gain, we used the difference between achievement at the start and at the end of the grade.

Simulated student results (of 16 simulations) in grade 1 and grade 4 were analyzed using multi-level analyses (Bryck & Raudenbush, 1992). Differences between schools and variables explaining the difference were located by using the general multi-level model (model A).

$$\textit{student level} \quad (\textit{learning gain})_{ij} = \beta_{0j} + r_{ij}$$

$$\textit{class level} \quad \beta_{0j} = \gamma_{00} + \sum_{h=1}^{n} \gamma_{0h} * X_h + u_{0j} \qquad \textit{(model A)}$$

Where:

$(\text{learning gain})_{ij}$	Learning gain of student i in class j
β_{0j}	Expected average learning for class j
r_{ij}	Unique effect for student i in class j
γ_{00}	Overall mean achievement
u_{0j}	Unique effect of class j on mean achievement conditioning for all other effects
γ_{0h}	Difference in mean achievement class variable h
X_h	Class variable h

The class variables (X_h) available for explaining the differences between classes are school type ($TYPE_j$), teacher time available (T_j), selection effect (K_j), average learning gain (μ_j), teachers' standard (S_j) and heterogeneity of achievement (σ_j). Of course average learning gain, teachers' standard and heterogeneity of achievement cannot be brought into the equation together, because standard is formulated as a function of average achievement and heterogeneity.

Results

Averaged results for grade 1 and grade 4 are summarized in Table 2.

The results show that the percentage of variation between classes in grade 4 is much higher than in grade 1. On average differences between schools account for about 25% of the variation in learning gain in grade 4, whereas in grade 1 only 6% of the variation in learning gain is due to differences between schools. This implies that learning gain becomes more dependent on class characteristics for the higher grades.

Table 2 Results from multi-level analyses

	Average number of students	Number of classes	Average learning gain (standard deviation)	Variance between classes (%)	Variance explained		Estimation of fixed effects		
					Between classes variance explained (%)	Total variance explained (%)	σ_j	K_j	T_j
grade 1	4750	148	2.4 (3.0)	5.8	76.4	4.4	0.19	-1.18	0.01
grade 4	3081	148	1.3 (1.7)	24.6	81.2	20.0	0.26	-1.39	0.01

K_j selection effects of class
σ_j standard deviation of learning gain of students in class
T_j teacher time available for class

For grade 1 the also classroom standard explains some variance. A surprise is the fact that the sign of the effect of the standard on learning gain is negative. This implies that lower average achievement or standards give higher learning gains. In general it is expected that these variables have a positive effect on learning.
Another reverse effect can be found for the selection effect K_j on learning gain. Hence students in schools which select higher achieving students, will make less learning gain.
Teachers' time, T_j, and classroom heterogeneity, σ_j, have a positive effect on learning gain made, although σ_j has a higher impact on learning in grade 4 than it has in grade 1. The possitive effect could be expected beforehand as the learning gain formula comprises linear multiplications factors of T_j and σ_j.

In empirical research on differences between secondary schools in the Netherlands, the between schools variance ranges from 7% to 40% of variation in student achievement (Roeleveld, 1987; Scheerens et al., 1989; Witziers, 1992; Luyten, 1994). Variables responsible for these differences are among others 'opportunity to learn', 'amount of homework', 'type of school' and 'teachers' experience'. The differences in simulated learning due to differences in classes lay in line with the results from empirical research, eventhough variation for grade 1 data rather low and for grade 4 the variation is rather high. Explanatory school variables from empirical reserach and explanatory classroom variables in the simulation can hardly be compared, since in research more descriptive variables are used for explaning school policy or school characteristics. The translation of these variables to classroom variables in the simulation cannot easily be made.

Experiment III Hypothetical Effects of Policy Changes

The main purpose of development of the simulation model was to be able to generate hypothetical effects for two major innovation for the Dutch secondary educational system. It was assumed that these innovations would change the standards set in the lower tracks and would improve achievement of students from low social background by increasing instruction time.
Thus four policy scenario's were derived from these innovations:

1. current situation;
2. higher standards in JVE and IGSE;
3. more time for low achieving students, and;
4. more time for students from low social backgrounds.

From the sensitivity analyses of standards and time from the previous sections it is expected that higher standards will increase achievement levels and that more time also results in higher achievement levels. The policy scenarios are all adjustments of teachers standard and allocation of teacher time and were implemented in the simulation model.

BEST COPY AVAILABLE

Technical aspects

For scenario 1 all variables have their default level. In scenario 2 higher standards are forced upon the lower tracks. Standards in HGSE and PUE are set on the default level, i.e. half a standard deviation above the average. Standards in JVE and IGSE are set 0.75 and 0.6 standard deviations above the average. In scenario 3 and 4 each student gets a minimum amount of time. For students performing below the standard time allocation in scenario 3 depends on the gap between standard and achievement. The same counts for scenario 4, but to with low SES even more time is allocated.

The differences in student throughput, generated by the policy scenarios are studied by using the theory of Markov Chains (Winston, 1987, p. 756). Education is seen as a system where students move with some probability from one track or grade to another one. At the end of their education the students reach some grade: drop-out, JVE or IGSE certificate, HGSE or PUE grade 5. Here their education ends.

Results

As shown in Table 3 average achievement rises when standards increase in JVE and IGSE. A slight decrease in achievement occurs when time is unequally distributed among students. Time is scarce and when more time is spent on low achieving students it means that high achieving students are given less time. Hence low achieving students will improve at the expense of high achieving students.

Table 3 Achievement levels for policy scenarios

	SES low		SES medium		SES high		Global	
Scenario	achievement	drop-out	achievement	drop-out	achievement	drop-out	achievement	drop-out
1	110.8	23%	115.8	12%	122.0	3%	116.5	12%
2	114.4	33%	117.5	18%	122.6	6%	118.3	18%
3	112.0	27%	115.5	13%	120.6	3%	116.2	13%
4	109.6	22%	111.0	12%	120.4	2%	115.0	11%

Next to a changing percentage of drop-outs due to policy changes, other student shifts occur for the policy scenario's.
Scenario 2 does also have an effect on students in the higher tracks as more HGSE students switch to lower track and less students switch to PUE. In PUE on the other hand, students have a higher probability of reaching the fifth grade.
Besides a decreasing performance level, scenario 3 and 4 also result in a diminishing number of students switching tracks and students are more likely to finish their track.

6 DISCUSSION

As there is a lack of theory in school effectiveness research, this paper presents a formal model of learning of a multi-level nature, which can be used to simulate student learning in a classroom setting. The variables causing learning gain were selected from theory and empirical research. Relations between variables were quantified using school organizations models, learning models and the model of De Vos (1989) to specify the relation between standard, grouping and learning.

The paper describes a validation experiment in which the simulated data are compared with data from a cohort study. In general correlations between student background and simulated achievement were of about the same sign and height as correlations within the cohort data. Because student background only has a minor effect on simulated learning, it is not clear whether these results imply a true relationship or whether these results are just a result of chance. The correlational stuctrure is weak: correlations range from between about plus and minus 0.20.

From the second experiment one can conclude that the model is able to generate differences between schools or classes. These differences range from 6% of total variation in grade 1 to 25% in grade 4. Variables explaining a large part of these differences are teacher time available and heterogeneity of achievement in the class. Standards only have a minor explaining value of differences between simulated learning. But in the simulation it are not the standard causing learning gain, but the gap between standard and achievement. So on a student level this gap could be able to explain differences.
In general it can be said that: "what goes into the model, also comes out". For example the large effect of time and heterogeneity which was found, is a result of time and heterogeneity which was put in as a multiplication factor into learning gain calculations.

Within the last experiment the model generated some hypothetical effects of policy scenario's. According to the model higher standards in the lower tracks generate higher performance levels, but also imply an increasing drop-out, especially for students of a low social background. Furthermore effects were found in the higher grade, when only changes in the lower grades are implemented.
Policies focussing on time distribution among students of differing social status, resulted in a decrease in drop-out numbers and a fall in performance. Additionaly the probability of switching tracks fell and more students stayed in their original tracks. This would imply that student get more equal chances.
The results found in this experiment were more or less expected beforehand, as higher standards were supposed to raise achievement and unequal time allocation were supposes to improve equality within education.

Besides the experiments described in the model, other experiment focussing on model validation and stability of results were executed. The outcomes showed that the model was less able to produce stable results regarding student movements. When comparing simulated student movements with real student movement, some significant differences were found. On the other hand results regarding performance and learning were reproducable and the datastructure was comparable with data from a national cohort. Hence the model needs to be adjusted especially in the domain of student movement.

Although the discussed model can be improved, we have shown how these models can be used to get more insight into student learning within a classroom environment. Furthermore, by experimenting with the model the effects of variable changes can be estimated by simulation. Hence, although relations can be interdependent and complex, simulation can make the net effects of variable changes clear.

LITERATURE

Bennett, S.N. (1978), "Recent research on teaching: A dream, a belief and a model", *British journal of educational Psychology*, 48, 127-147.

Bloom, B.S. (1976), *Human characteristics and school learning*, McGraw-Hill, New-York.

Bosker, R.J. & Guldemond, H. (1994), *A hierarchical simulation model to study educational interventions*, University of Twente, Enschede.

Bryck, A.S. & Raudenbush, S.W. (1992), *Hierarchical linear models: Applications and data analyses methods*, SAGE Publications, Newbury Park.

Carroll, J.B. (1963), "A model of school learning", *Teachers College Record*, 64, 723-733.

Cooley, W.W. & Leinhardt, G. (1975), *The application of a model for investigating classroom processes*, University of Pittsburgh, Pittsburgh.

Forrester, J. (1971), *World dynamics*, Wright-Allen Press, Cambridge, Mass.

Geurts, J. & Vennix, J. eds. (1989), *Explorations in policy analyses (Dutch: Verkenningen in beleidsanalyse)*, Kerkebosch, Zeist.

Geurts, J. (1983), "Social planning, system thinking and simulation (Sociale planning, systeemdenken en simulatie)", *Simulation and Social Systems (Dutch: Simulatie en sociale systemen)*, 3, 2/3, 31-54.

Haertel, G.D., Walberg, H.J. and Weinstein, T. (1983) "Psychological models of educational performance: A theoretical synthesis of constructs", *Review of Educational Research*, 53, 75-91.

Harnischfeger, A. & Wiley, D.E. (1976), "The teaching learning process in elementary schools: A synoptic view", *Curriculum Inquiry*, 6, 5-43.

Leinhardt, G. (1980), "Modelling an Measuring Educational Treatment in Evaluation", *Review of educational research*, 50, 393-420.

Levin, G. & Roberts, E.B., (1976), *The dynamics of human service delivery*, Ballinger Publishing Co., Cambridge

Luyten, J.W. (1994), *School effects: stability and malleability*, University of Twente, Enschede (dissertation).

Roeleveld, J. (1987), "Differences in outcomes of Dutch secondary schools (Dutch: Verschillen in resultaten tussen Nederlandse scholen voor voortgezet onderwijs)", In: *J. Scheerens & W.G.R. Stoel (eds.), Effectiveness of educational organizations (Dutch: Effectiviteit van onderwijsorganisaties)*, Swets & Zeitlinger, Lisse.

Scheerens, J. (1992), *Effective schooling*, Cassel, London.

Scheerens, J., Nanninga, H.C.R. & Pelgrum, W.J. (1989), "Generalizaibility of instructional and school effectiveness indicators across nations, *International journal of educational research*, 13

Snow, R.E. (1973), "Theory construction for research on teaching", In: *R.M.W. Travers (ed.), Second handbook of research on teaching*, Rand McNally, Chicago.

Vos, H. de (1989), "A rational-choice explanation of composition effects in educational Research", *Rationality and Society*, 1, 2, 220-239.

Winston, W.L. (1987), *Operations research: applications and algorithms*, Duxbury Press, Boston.

Witziers, B. (1992), *Coordination in secondary schools (Dutch: Coordinatie binnen scholen voor voortgezet onderwijs)*, University of Twente, Enschede (dissertation).

APPENDIX I Secondary Education in the Netherlands

The Dutch educational system is divided into three main categories:
- primary education for the age of pupils from 4 through 12
- secondary education, duration 4 to 6 years
- senior vocational education and higher education

The system for secondary education is categorical. At the age of thirteen pupils have to decide what type of school is most suited to the individual pupils' aims and cognitive capacities. Generally speaking they have four options (ranked in increasing difficulties);
1. Junior Vocational Education (JVE), duration 4 years;
2. Intermediate General Secondary Education (IGSE), duration 4 years;
3. Higher General Secondary Education (HGSE), duration 5 years, and;
4. Pre-University Education (PUE), duration 6 years.

A general scheme of student streams is given in Figure 5. Generally speaking upward switches between the school types are possible after acquiring a certificate. Down in the figure but happens for example instead of repeating a grade for the second time. The grades within the simulation model are indicated by the rounded box.

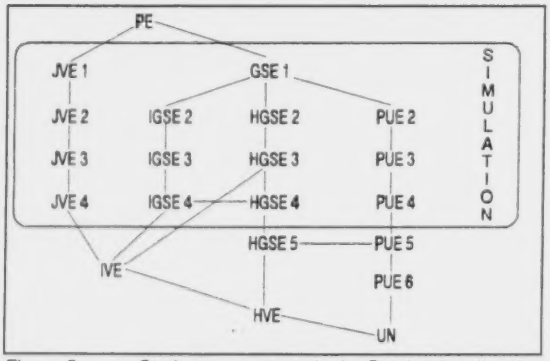

Figure 5 Student movements in the Dutch educational system

HGSE	Higher General Secondary Education (5 years)
HVE	Higher Vocational Education (4 years)
IGSE	Intermediate General Secondary Education (4 years)
IVE	Intermediate Vocational Education (4 years)
JVE	Junior Vocational Education (4 years)
PE	Primary education (6 years)
PUE	Pre-university Education (6 years)
UN	University (4 years)

APPENDIX II Mathematical formulation of the model

The following set of mathematical equations is used to determine learning and track placement in the simulation model.

Selection Effects $A_{ij} = \hat{A}_{ij} + \psi_{jk} + \psi_k$

where $\psi_{jk} \sim N(0, 0.10\sigma_j)$ $\psi_k \sim N(0, 0.05\sigma_j)$

Teachers' Standard $S_j = \mu_j + \kappa\,\sigma_j$

Expected Achievement $A_{ij} = \beta_0 + \beta_1\, SES_{ij} + \beta_2\, IQ_{ij} + \beta_3\, RATING_{ij} + \beta_4\, SEX_{ij} + e_{ij}$

Teachers' Time $\tau_{ij} = \dfrac{\log_{10} T_{ij}}{\log_{10} 55}$

where $T_{ij} \sim U(10, 100)$

Learning Gain

$\Delta_{ij} = u\,(1-Q)\,\tau_{ij}\,\sigma_j$ *when* $e_{ij} < 0$ $A_{ij} < S_j$

$\Delta_{ij} = u\,(1-Q)\,\tau_{ij}\,\sigma_j\,\xi_1$ *when* $e_{ij} \geq 0$ $A_{ij} \geq S_j$

$\Delta_{ij} = u\,(1-Q)\,\tau_{ij}\,\sigma_j\,\xi_2$ *when* $e_{ij} < 0$ $A_{ij} \geq S_j$ *where* $u \sim U(0,1)$

$\Delta_{ij} = (u-\frac{1}{2})\,\tau_{ij}\,\sigma_j$ *when* $e_{ij} \geq 0$ $A_{ij} \geq S_j$

Track Placement

move up when $S_j - \delta_1\,\sigma_j \leq A_{ij} \leq S_j + \delta_2\,\sigma_j$

repeat when $A_{ij} < S_j - \delta_1\,\sigma_j$

switch up when $A_{ij} > S_j + \delta_2\,\sigma_j$

Selection Effects

A_{ij}	Achievement corrected for selection effects	Differences between schools account for 10 percent of the total variation in initial achievement. Generation effects within schools are 5 percents (Bosker & Guldemond, 1990). The simulation model incorporates these effects by a random allocation of children to classes and then applying a correction for the selection effects. Achievement levels are artificially raised, so it looks like students are non-randomly allocated to classes and schools select their pupils. ψ_{jk} and ψ_k are randomly determined from a normal distribution with mean 0 and standard deviation $0.10^*\sigma_j$ and $0.05^*\sigma_j$ respectively.
$\hat{A}_{ij}$	Achievement not corrected for selection effects	
σ_j	Standard deviation of achievement	
ψ_{jk}	Selection effect for class j within school k	
ψ_k	Selection effect for school k	

Teachers' Standard

S_j	Standard	The standard can be seen as some positively motivating striving point, if a student sees other classmates functioning near the standard. When the standard is set to high, it may work demotivating. Therefore the standard must be somewhere above the average achievement level of a class, depending on the heterogeneity of the class. Default value for κ is 0.5.
μ_j	Average achievement	
σ_j	Standard deviation of achievement	
κ	Effect of standard deviation on standard	

Expected Achievement

A_{ij}	Achievement level of student	The degree of overachievement is measured by comparing a pupil's actual performance level with his predicted performance level. This prediction is based on socio-economic status, IQ, gender and teachers rating at the end of primary education. The regression equation relates achievement of each individual student to background variables and parameters are estimated for a certain classroom composition (references). The regression residual, e_{ij}, reflects de degree of overachievement: a negative residual means the student is an underachiever, whereas a positive residual reflects the overachievement of a student
β_h	Regression parameters, h = 0,...,4	
SES_{ij}	Social economic status of student	
IQ_{ij}	Students' intelligence	
$RATE_{ij}$	Teachers rating at the end of primary education	
SEX_{ij}	Gender of student	
e_{ij}	Regression residual	

BEST COPY AVAILABLE

Teachers' Time

τ_{ij}	Effect of amount of teacher time on learning	The time-ratio τ_{ij} is a transformation of the time a teacher in class j uses effectively for student i (T_{ij}). Hence T_{ij} is a combination of the time a teacher has available for the students and the how effectively the teacher uses this time. This effective instruction time is randomly allocated to the teacher at the start of the year and is equally distributed between 10 and 100%, with an average of 55%.
T_{ij}	Teacher time available	It is reasonable to expect that the effect of extra time on learning gain declines, thus marginal learning gain decreases as time decreases. This means that a time increase from 10% to 20% has more effect on learning gain than a rise in time from 80% to 90%. This non-linear increase is modelled using logarithms.

Learning Gain

Δ_{ij}	Learning gain of student i	For each student, learning in a year is determined in one step. For the purpose of learning gain calculations a distinction is made between underachievers and overachievers with respect to their learning gain possibilities. Furthermore both groups of students can achieve below or above the standard, thus implying four learning calculation strategies: i. Overachievers achieving above standard make no learning gain on average. Their learning gain is uniformly distributed around 0. ii. For overachievers achieving below the standard it is assumed that some learning gain can be made, because the standard is not yet reached and may motivate the student. iii. Underachievers achieving above standard are somewhat demotivated, but are able to make some learning gain, because they achieve below expectation and hence are able to perform better. iv. Underachievers achieving below standard, can make the highest gains. The values for ξ_1 and ξ_2 are 1 and 2 respectively.
u	Random correction for variables not in the model between 0 and 1	
1-Q	Non-linear effect of standard on learning	
τ_{ij}	Effect of amount of teacher time on learning	
σ_j	Standard deviation of achievement	
ξ_1, ξ_2	Correction for overachievement respectively achieving above standard	

Track Placement

A_{ij}	Achievement level of a student	Track placement decisions are made at the end of each year. Such a decision depends upon the level on which a student performs and some critical value based on the standard the teacher has set for the class and classroom heterogeneity. The following movements can be distinguished: 1. moving up to the following grade; 2. switching to the same class of a higher type of education; 3. dropping to a lower type of education, higher class (this occurs when students otherwise have to repeat the grade for the second year); 4. repeat a grade; 5. drop out (only from the lowest track)
S_j	Teachers' standard	
σ_j	Standard deviation of achievement	
δ_1, δ_2	Effect of σ_j on the critical level for repeating respectively an upward switch	

BEST COPY AVAILABLE